ABSTRACT

A method for fast and efficient record retrieval in large databases using cyclical redundancy check (CRC) computations as hash functions. Two hash values are computed for each record's key using the CRC-CCITT and CRC-16 generator polynomials. The two CRC values then are combined into a four-byte composite hash value that represents a binary signature of the record's key. Alternately, a single CRC-32 value can be used as a four-byte hash value. In most cases, this four-byte hash value uniquely identifies the record's key. An index file is constructed using a hybrid search method, part hash table and part linear search. The index file is searched to find a match for the four-byte hash value and the record's offset is obtained. The record's offset is used to retrieve the record from the database.